



TITLE:

道徳と言語のアナロジー説の批判的検討 : 感情説との比較を通じて

AUTHOR(S):

田中, 泉吏; 中尾, 央

CITATION:

田中, 泉吏 ...[et al]. 道徳と言語のアナロジー説の批判的検討 : 感情説との比較を通じて. 科学哲学科学史研究 2009, 3: 1-19

ISSUE DATE:

2009-02-28

URL:

<https://doi.org/10.14989/72811>

RIGHT:

道徳と言語のアナロジー説の批判的検討

感情説との比較を通じて

A critical examination of the linguistic analogy of morality:
Through a comparison with emotional theory of morality

田中 泉吏*・中尾 央†

Senji TANAKA and Hisashi NAKAO

abstract

The aim of this paper is to critically examine the theory of linguistic analogy of morality by Mikhail, Harman, Hauser, and Dwyer through a comparison with emotional theory of morality by Haidt, Greene, and Prinz. The theory of linguistic analogy argues that moral judgments come from our unconscious moral grammar. Dual process theory by Greene, on the other hand, argues that moral judgments are based on our unconscious emotional response and conscious reasoning. We argue that the action analysis model in the theory of linguistic analogy may be able to reveal an important aspect of moral judgment, but that the theory of linguistic analogy should not underestimate the role of emotion in moral judgment.

§1 導入

20 世紀の後半に至るまで、道徳心理学ではピアジェやコールバーグの強い影響力の下で、道徳判断は意識的な推論によって生み出されると考えられて、感情が果たす役割は軽視されていた。しかし、20 世紀の終わりになって、神経心理学や臨床心理学 (Damasio 1994), また社会心理学や経済学 (Frank 1988) の中で再び感情に注目が集まるようになった。さらに fMRI などの、脳の働きを直接画像で見ることのできる技術が開発されたことも相俟って、道徳心理学においても感情の役割が再考されるようになってきている。このように、意識的な推論よりもむしろ感情が道徳判断を生み出す主要な要因であることを強調する立場を道徳の「感情説」と呼ぶ。この立場は現在、

* 京都大学大学院文学研究科 研修員 senji@fd6.so-net.ne.jp

† 京都大学大学院文学研究科 博士後期課程 hisashinakao@gmail.com

ハイト、グリーン、プリンツなど (Haidt 2001; Greene et al. 2001; Prinz 2008a) によって支持されている¹。

他方で、道徳を言語と類比的に捉え、チョムスキー派の理論言語学の枠組みに基づいて道徳の生得性を主張する「道徳と言語のアナロジー説」も盛んに論じられている。この立場は現在、ハーマン、ミハイル、ドゥワイア、ハウザーらによって支持されている (Harman 2000; Mikhail 2000; Dwyer 2006; Hauser 2006)。彼らによれば、道徳判断は言語文法にいくつかの重要な点で類似する「道徳文法」と呼ばれる²認知能力によって導かれるのだという。

感情説と道徳と言語のアナロジー説は両者とも、健常者に様々な道徳的ジレンマに対して回答してもらうという実験を利用しているという点で共通している（ときには全く同じジレンマを用いている場合もある）のだが、そこから導かれる結論は、両者の間で非常に大きく異なっている。そこで本稿では、両者の比較を通じ、道徳と言語のアナロジー説を批判的に検討することを試みる。以下ではまず、感情説の主張を（2節）、そして道徳と言語のアナロジー説の主張を分析した上で（3節）、この立場に対して挙げられている批判をいくつか取り上げて検討する（4節）。そして以上の分析を踏まえながら、両仮説を比較して今後の展望を探る（5節）。

§2 道徳の感情説

道徳の感情説とは、道徳判断を生み出す主要な要因は感情であるという主張である。例えばハイトは、健常な被験者に道徳的ジレンマに回答してもらったときに、その多くがすばやく道徳判断を下すのだが、判断の根拠について尋ねられると（明確に、あるいは何も）答えられないということに注目した (Haidt 2001)。彼はこのことに基づいて、道徳判断の生成においては意識的推論ではなく感情が支配的な役割を果たしているのだと結論している。

上記のようなハイトの結論は、グリーンらの神経科学実験 (Greene et al. 2001) によって支持されている。グリーンらは、健常者が様々な道徳的ジレンマに回答する際の脳の働きを fMRI を用いて調べた。その結果、被験者の反応とそのときの脳の活動

¹ 彼らは感情に注目するという点で共通しているが、道徳感情が生得的かどうかという点に関しては意見を異にする。ハイトとグリーンは生得説、プリンツは反生得説の立場に立つ。

² 道徳能力 (moral competence) や I-道徳性などとも呼ばれるが、本稿では一貫して「道徳文法」と呼ぶことにする。

が、「トロリー・ジレンマ」に代表されるジレンマ群と、「歩道橋ジレンマ」に代表されるジレンマ群とで、大きく異なることが明らかになった。以下では、説明を簡単にするために、トロリー・ジレンマと歩道橋ジレンマの2つのみを考えることにするが、これらに類似するジレンマでも同様の実験結果が出ているということに留意されたい。

「トロリー・ジレンマ」においては、暴走したトロリーが線路上の5人の人に向かって進んでおり、このままでは彼らがトロリーに轢き殺されてしまう。彼らを救う唯一の手立ては線路の切り替えスイッチを押すことだが、切り替えた先の線路にも1人の人がおり、スイッチを押せばその人が轢き殺されてしまう。以上の状況が説明された後で、被験者は「あなたが5人を救うためにスイッチを押すことは適切か」と問われ、回答を求められた。実験では、ほとんどの被験者が「適切である」と答えた。

「歩道橋ジレンマ」においては、暴走したトロリーが線路上の5人の人に向かって進んでおり、このままでは彼らがトロリーに轢き殺されてしまう。あなたは、線路の上にかかっている歩道橋の上で見知らぬ大男の横に立っている。このとき、5人の人を救う唯一の手立ては見知らぬ大男を歩道橋から線路の上に突き落とすことである。そうすれば男は死ぬだろうが、トロリーは男に衝突するため5人は助かるだろう。以上の説明がされた後で、被験者は「あなたがこの見知らぬ男を突き落として5人を救うことは適切か」と問われ、回答を求められた。実験では、ほとんどの被験者が「適切ではない」と答えた。

いずれのジレンマでも1人を犠牲にして5人を救うという点では同じなのに、どうして両ジレンマでは人々の判断が異なるのか。それを説明するため、グリーンは次のような仮説をたてた。歩道橋ジレンマでは、トロリー・ジレンマの場合よりも感情が誘発され、それが人々の道徳判断の違いを生み出したのではないか。そこでグリーンは、その他の道徳的ジレンマでも同様の結果が出るかどうかを調べるために、道徳的ジレンマを「個人的」ジレンマと「非個人的」ジレンマに分類して実験を行った。個人的なジレンマとは、(a) 特定の人や特定の集団の成員に対して、(b) 深刻な身体損傷をもたらすと予測され、(c) その損傷が現存する脅威をただ逸らしただけの結果ではない場合、である。これらの条件を一つでも満たさない場合、それは非個人的なジレンマと呼ばれる。これらの条件は、(a) 傷つける相手、(b) 傷つけたという事実、(c) 誰が傷つけたのか、を明確にするための条件である。歩道橋ジレンマはこれらの条件をすべて満たすため個人的な道徳的ジレンマであるが、トロリー・ジレンマでは現存する脅威（暴走するトロリー）をただ逸らす（スイッチを切り替える）だけなので、(c) の条件を満たさない。したがって、トロリー・ジレンマは非個人的な道徳的ジレン

マである。

以上のような仮説を検証するため、彼は fMRI を用いてジレンマに直面した被験者の脳の反応を測定した。そして実験の結果、トロリー・ジレンマなどの非個人的な道徳的ジレンマでは、高次の認知に関する部位（下頭頂葉や中前頭回）が個人的な道徳的ジレンマの場合よりも（また、非道徳的なジレンマの場合よりも）有意に活性化していた。これらの部位はワーキングメモリと結びついていることが知られている。そして歩道橋ジレンマなどの個人的ジレンマでは感情に関する部位（角回や後帯状回、内側前頭回）が反応しており、非個人的ジレンマの場合よりも（また、非道徳的ジレンマの場合よりも）有意に活性化していた。グリーンらの仮説は神経科学の観点から一定の支持を受けたわけである。

このように、グリーンの研究は基本的に感情説の路線を採るわけだが、だからといって彼は意識的推論が道徳判断において全く役割を果たさないとやっているわけではない。彼はこのことを、個人的な道徳的ジレンマでも被験者の反応に違いが見られるものがあることに注目して述べている（Greene et al. 2004）。彼が注目したのは「泣き叫ぶ赤ん坊のジレンマ」と「嬰兒殺しのジレンマ」である。両ジレンマとも個人的なジレンマであり、自分の赤ん坊を死に至らしめるという点で共通しているが、被験者の反応は異なっている。では、どうしてこのような差異が見られるのだろうか。前者のジレンマでは、あなたの村が敵国によって占領され、敵国の兵士は残った村人を皆殺しにしようとしている。あなたと村人が隠れているときに、あなたの赤ん坊が大声で泣き出した。あなたはとっさに赤ん坊の口を手で塞いだが、そのままでは赤ん坊は窒息死してしまう。しかし、その手を離せば赤ん坊の泣き声のために敵国の兵士に見つかり、あなたとあなたの子どもを含めた村人全員が皆殺しにされてしまう。このとき、被験者は「あなた自身と村人を救うために、あなたの赤ん坊を窒息死させることは適切か」と尋ねられる。後者のジレンマでは、あなたは妊娠した15歳の女の子である。あなたは何とか妊娠を隠し通し、ある日ロッカールームでこっそり出産した。あなたには産まれた赤ん坊を育てることなどできないので、出産したことを隠して赤ん坊をゴミ箱に棄てられればどんなに楽かと考える。このとき、被験者は「あなたが楽な人生を送るために、あなたの赤ん坊をゴミ箱に棄てることは適切か」と尋ねられる。

これらのジレンマは両者とも(a)～(c)の条件を満たす個人的なジレンマである。また、両者とも自分の赤ん坊を死に至らしめる点で共通している。しかし、両ジレンマで子供を殺すことが許されるかどうかを被験者に尋ねたところ、被験者の反応は大きく異なった。前者では比較的時間をかけた後（平均で8秒強後）に判断が下された

ものの、被験者間で判断は一致しなかった。他方、後者では「赤ん坊を殺してはいけない」という判断が、多くの人によって比較的素早く（平均で3秒弱後）下された。この際の脳反応をfMRIで測定したところ、嬰兒殺しのジレンマは歩道橋ジレンマと同様に感情に関わる部位がより活性化したが、泣き叫ぶ赤ん坊のジレンマでは、感情に関わる部位だけではなく、前帯状回皮質や背外側前頭前皮質といった認知制御や推論に関わる部位が、より活性化した。これらの結果をグリーンらは次のように解釈している。嬰兒殺しのジレンマでは、感情が「赤ん坊を殺してはならない」という道徳判断を導いた。泣き叫ぶ赤ん坊のジレンマは、自分の赤ん坊を殺すべきかどうかを判断するという点で嬰兒殺しのジレンマと同じだが、感情に従って「赤ん坊を殺すべきではない」と判断することが合理的な結果をもたらさないため、認知的推論が動員された。そしてその結果、判断までに時間がかかり、被験者によって意見が分かれたのである。

以上のように、グリーンは道徳判断の生成における推論の働きについても、考察を加えている。すなわち、我々が合理的（適応的）な道徳判断を下すことができるのは、感情と推論という二種類の心理過程を持っているからである。彼は、このような自身の立場を「道徳判断の二重過程モデル」と呼んでいる。

§3 道徳と言語のアナロジー説

前節で検討した感情説は道徳判断における感情の役割を強調していたが、これに対して、道徳判断をあくまでも認知計算過程として特徴付ける立場がある。この立場は、チョムスキー言語学の枠組みに基づいて、人間の道徳は言語文法（生成文法）に類比的なもの、すなわち「道徳文法」として理解できると主張する。この立場は、「道徳と言語のアナロジー説」と呼ばれている（Cf. Mikhail 2000, 2007; Dwyer 1999, 2006; Harman 2000; Hauser 2006; Hauser et al. 2007, 2008）³。

道徳と言語のアナロジー説が答えるべき課題は、以下の5つに分けられる（Mikhail 2007, p. 144）。

（1）道徳文法とは何か

³ ミハイルとハーマンは、道徳と言語のアナロジー説の着想をジョン・ロールズ『正義論』第9節から得ており、自分達の研究プログラムをロールズの構想を展開したものと述べている。したがって、彼らにとっては、道徳文法や普遍道徳文法の理論は規範倫理学の理論と密接に関係することになるだろう。こうした論点は非常に興味深いが、本稿のテーマはロールズ解釈や道徳心理学と規範倫理学の関係にはないので、これに関する考察は別の機会に譲ることとしたい。

- (2) 道徳文法はどのようにして獲得されるか
- (3) 道徳文法は実際にどのように使用されているか
- (4) 道徳文法は脳内でどのように実装されているか
- (5) 道徳文法はどのように進化してきたか

道徳と言語のアナロジー説に沿った研究が展開されてからまだ十年も経っていないため、具体的なモデルが提出されているのは(1)(2)に関してだけである。したがって、本稿ではこれらに関してのみ論じ、(3)(4)(5)については扱わない。

本節ではまず、言語のアナロジー説の「道徳文法が存在する」という主張と、その論拠について述べる(3.1節)。次に、課題(1)に対する回答となる、道徳文法の「行為分析モデル」について説明する(3.2節)。さらに、課題(2)に対する道徳と言語のアナロジー説の回答である、人間には生まれつき「普遍道徳文法」が備わっているという主張について検討する(3.3節)。そして、普遍道徳文法の「原理・パラメータ・モデル」について考察する(3.4節)。最後に、感情説との比較によって、両者の相違点と共通点を明らかにする(3.5節)。

3.1 道徳文法

道徳と言語のアナロジー説では、すべての正常な人間は道徳に特化した認知計算過程(モジュール)を持ち、我々の道徳判断はこれによって可能になっていると主張される。この認知計算過程は、チョムスキー言語学における「言語文法」にいくつかの重要な点で類似すると考えられており、それゆえ「道徳文法」と呼ばれる。ここでいう「文法」とは、すべての正常な人間が持つ認知システム(生物器官)のことであって、学校で教えられる国語や英語の文法とは意味が全く異なることに注意されたい。

道徳文法が存在するという論証は、言語文法が存在するという論証とのアナロジーによって説明される。後者の論証は、以下のようなものである。人間が観察することのできる文は有限である。それにもかかわらず、正常な大人であれば、潜在的に無限の数の文を作り出すことができる。また母語であれば、それまで一切聞いたことがないような表現についても、それが統語論的に正しいかどうかということや、意味が曖昧であるかどうかということについて、直観的な判断を下すことができる⁴。つまり、人

⁴ 例えば我々は、「彼女の妻が赤い知識を眺めている」という表現が、意味は支離滅裂だが統語論的には正しいと直ちに理解できる。また、「美しい黒髪の女の子」という表現の意味が曖昧である(美しいのは髪か女の子か)ということも直観的に理解できる。

間はいったん母語を完全に身に付けてしまえば、その言語表現について様々な判断を潜在的に無限に下すことができるようになる。しかし、我々の脳の記憶力は有限なので、無限の表現の知識を蓄えておくことはできない。したがって、我々の脳には有限の語句リストから無限の表現を作り出し、それについて判断するための、ある種のレシビまたはプログラムがなければならない。これが「言語文法」である (Mikhail 2000, pp. 52-54)。

このように、言語文法の存在論証はアブダクション (最善の説明への推論) の形式をとる。すなわち、我々が観察する様々な言語現象の最善の説明は、言語文法が正常な言語使用者の心/脳に存在するということであり、この仮定なくしては、観察される現象はほとんど説明不可能である、というのである (ibid., p. 51)。

道徳文法の存在論証も同様の仕方で行われる。人間は誰も有限の数の道徳判断しか観察することができない。それにもかかわらず、我々は新奇な道徳判断を潜在的に無限に下すことができる。また、それまで出合ったことのない全く新しい状況に対しても道徳判断を下すことができる。しかし我々の脳の記憶力は有限なので、無限の道徳判断を蓄えておくことはできない。したがって、言語のアナロジー説によれば、様々な道徳判断を無限に下すための何らかの認知システムが我々の心/脳に備わっていると仮定しなければ、我々が日常的に行う道徳判断はほとんど説明できない。すなわち、我々の心/脳には「道徳文法」という認知システムが存在するのだと言う (ibid., pp. 54-57)。この論証は、どのような文化にも何らかの道徳体系が存在するという事実によっても支持される。実際、どの自然言語にも基本的な義務概念を表現する語句が存在している (Mikhail 2007, p. 144)。

3.2 行為分析モデル

では、言語文法と類似した道徳文法が存在するとすれば、それは具体的にどのような認知システムと考えられるのだろうか。ハウザーとミハイルは、言語とのアナロジーにより、道徳文法は複雑な規則体系であっていくつかのサブモジュールから構成されると考えている。すなわち、ある行為が刺激として入力されて道徳判断を出力する際、この入力と出力の間をつなぐサブモジュール群が想定されているのである。これらのサブモジュール群は、行為を一定の規則に従って分析し、最終的に道徳判断を導くとされる。このような道徳文法のモデルは「行為分析モデル」と呼ばれる。

ミハイルによれば、道徳文法には少なくとも「転換規則」、「構造記述」、「義務論規

則」という3つの要素がなければならない(Mikhail 2007)。刺激入力はこの順序で処理されていき、最終的に道德判断を出力する。転換規則とは、刺激を無意識的な心的表象に転換するためのサブモジュールである。このような転換規則の想定は、認知科学の理論では一般的なものであり、例えば視覚や言語の理論にもみられる。すなわち我々は、網膜上の二次元の刺激から三次元の表象を復元することができるし、切れ目のない聴覚刺激パターンから個別の単語を認識することができる。認知科学では、生得的なモジュールが入力刺激に欠如している情報を復元する役割を果たすと想定されている。道德と言語のアナロジー説においても、同様の仕方ですubモジュールが想定されるのである。

では、転換規則は何を復元するのだろうか。ある行為という刺激の中に、その行為の手段や結果、あるいは行為に含まれる傷害(battery)という要素が直接的に明示されているわけではない。何が手段で何が結果かというようなことは、我々の心的表象内において初めて認められる性質である。転換規則はこれらの性質を復元し、それらを含む行為の十分な構造記述を可能にするのである。

この転換規則自体も、複数のサブモジュールから構成される規則群であると考えられている。具体的には、次の5つの規則が考えられている。一つ目は、行為の特定の要素を適切な時間順序に並べる規則である。二つ目は、それらの要素間の因果関係を明らかにする規則である。三つ目は、特定の結果に「よい」、「悪い」という性質を付与する規則である。四つ目は、行為者がよい結果を意図して悪い結果を避けると仮定することによって、結果と副産物を区別する規則である。五つ目は、傷害の表象を導く規則である。以上の5つの規則を適用することによって得られる心的表象が、構造記述である。構造記述は図1のような「行為樹」として表現される。

義務論規則とは、この構造記述から特定の道德判断を導き出すための規則である。行為に傷害の要素が含まれるとき、それが副産物の枝部分にある場合は許されるが、手段の幹上にある場合は許されないというような規則が、その一例と考えられている(ibid., pp. 145–148)⁵。

例えば、トロリー・ジレンマの行為樹では傷害が副産物であるが、歩道橋ジレンマの行為樹では傷害が手段となっている。ミハイルによれば、このことがトロリー・ジレンマと歩道橋ジレンマに対する人々の道德判断の違いを説明するのであって、グリー

⁵ 既に述べたように、この義務論規則は道德文法というモジュールを構成する一つのサブモジュールであって、我々が日常的に従っているような道德規範とは別物である。

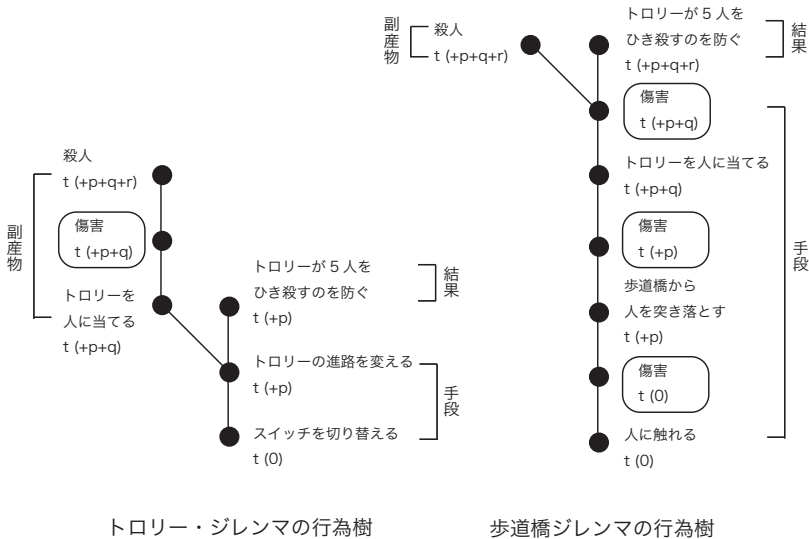


図1 Mikhail 2007, p. 150 より。

ンが言うようにジレンマが「個人的」であるか「非個人的」であるかは本質的な区別ではないと主張する。

3.3 普遍道徳文法

以上のような一連の無意識的な処理過程を経て道徳判断を導くのが、道徳文法モジュールである。このように複雑かつ高度な認知計算を行うメカニズムがいかにして発達過程において獲得されるかは、言語文法の場合と同様、重要な課題であろう。言語文法の場合は、生得的基盤（普遍文法）によって獲得が可能になっていると考えられている。道徳と言語のアナロジー説もまた、道徳でも同様な生得的基盤（普遍道徳文法）が存在し、これにより道徳文法の獲得が可能になっていると考えられている。言い換えれば、道徳文法の性質と発達パターンには一定の遺伝的制約が存在するということである。

普遍道徳文法を想定する根拠となるのが、いわゆる「刺激の欠如に基づく議論」である（Dwyer 2006, pp. 239–250）。これは普遍文法を支持する議論としても使われていたが、それによれば、幼児期の言語習得は十分な環境刺激がなくても達成される。このように十分な経験がないにもかかわらず言語文法を身に付けられるのであれば、

何らかの生得的な要素，すなわち普遍文法が存在しているに違いないということになる．同様に，道德判断は発達極めて早い時期(2,3歳くらい)で可能になる．このように，道德習得に必要とされる十分な経験がない段階で道德文法を部分的にでも身に付けられるのであれば，何らかの生得的な要素，すなわち普遍道德文法が存在しているに違いない，というのである(Dwyer 1999; Mikhail 2007)．

普遍道德文法は，あらゆる地域・社会の文化で道德が見られるという道德の普遍性を説明する上で極めて有効である．しかし，道德には普遍性だけでなく多様性も存在する．この多様性と普遍性を併せ持つという特徴は，言語にも見られる．言語の普遍性とは，どの言語の文法も生成文法理論に従った構造を持っているということである．言語の多様性とは，例えば英語と日本語では語順が異なることや，日本語やイタリア語では無主語でも適格な文とされるが，英語では無主語の文は不適格とされるといった差異である．道德の場合も，殺人・危害が禁止される対象や，援助・分配・交換が推進される対象，インセスト・タブーの適用対象などが文化ごとに大きく異なるということが，人類学によって明らかになっている(Hauser et al. 2008)．

言語がこの普遍性と多様性という一見相反する特徴を持ち合わせているということは，1970年代までの生成文法理論にとって一つの大きな問題であった．しかし，この問題は1980年代以降にチョムスキーが「原理・パラメータ・モデル」を発展させたことによって解消されたと言われる(田窪ほか1998)．そして，道德と言語のアナロジー説においても同様に原理・パラメータ・モデルが導入されて，道德の多様性も説明することが可能になるというのだ．

3.4 原理・パラメータ・モデル

言語の原理・パラメータ・モデルによれば，人間の言語文法はいくつかの原理とそれに伴うパラメータによって構成される．原理とは一定の文構造のみを認める制約のことであり，どの言語もこの制約に従っている．原理にはパラメータが付属しており，これにより文化ごとで制約の中に多様性が出現する．パラメータの具体例としては，語順や主語の有無などが挙げられている．このモデルによれば，人間は皆パラメータ値が未設定の原理の体系を持って生まれる．そして，環境によって話されている言語が異なるので，パラメータ値が環境によって異なる値に設定される．そうすると，生得的な普遍文法と成人が持つ言語文法は同じ認知能力の初期状態と最終状態を指し示す言葉ということになる．

道徳と言語のアナロジー説によれば、道徳文法も言語文法と同様にいくつかの原理と、それに伴うパラメータからなる。道徳の「原理」の候補としては、「G を殺すことは悪である」というような規則が考えられている。この場合、パラメータは G であり、これには特定の集団（もしくは個人）が当てはまる。このパラメータは社会や個人によって異なったものになりうる。例えば、現代の民主主義社会では G には人類が当てはまるが、ナチス政権化のドイツのような国家では特定の民族や支配階級しか当てはまらない。また、死刑制度や人工妊娠中絶の問題についての道徳的主張の不一致は、G の値の違いを反映しているのかもしれない（Cf. Harman 2000, p. 225）。そして、パラメータ値が確定している原理の体系が、個々人が持つ道徳文法ということになる。裏を返せば、パラメータ値が設定されていない原理の体系が、普遍道徳文法である。そうすると、普遍道徳文法と道徳文法は実は同じ認知能力の発達の初期状態と最終状態を指し示す用語ということになる。

このように、道徳と言語のアナロジー説によれば、普遍道徳文法は道徳文法の獲得を可能にさせる遺伝的発達プログラムである。しかし、だからといってその発達に環境刺激が不要だというわけではない。というのも、発達過程の環境刺激（経験）はパラメータ設定に必要なからである。そして普遍道徳文法と適切な環境刺激があれば、道徳文法を獲得することができるようになる。

以上のような道徳の原理・パラメータ・モデルは、道徳規範の普遍性と多様性の両方を説明することができる理論である。道徳規範に普遍性がみられるのは、人類が普遍道徳文法を共有するからである。また多様性がみられるのは、異なる環境下ではパラメータが異なる値に設定されるからである。このように、道徳の原理・パラメータ・モデルは非常に複雑かつ体系的な理論であり、生得説はその一部をなす主張である。そして、この原理・パラメータ・モデルは普遍道徳文法という生得的要素を認めつつも、パラメータ設定として環境要因が組み込まれているのである。

3.5 道徳と言語のアナロジー説と感情説の比較：共通点と相違点

ここで一度、本節で述べてきた道徳と言語のアナロジー説と 2 節で検討した感情説の共通点と相違点を明確にしておこう。

両者の第一の共通点は、道徳判断が直観的・無意識的に行われると主張する点である。したがってこの点においては、両者は共にピアジェ＝コールバーグ流の研究パラダイムに対して共同戦線を張っていると言うことができる。またこの主張は、感情説

の実験研究でも、道徳と言語のアナロジー説の実験研究でも、明らかにされた事実である。既に述べたように、感情説の立場に立つハイトとグリーン、そして道徳と言語のアナロジー説の強力な推進者であるハウザーやミハイルによれば、実験の中で被験者に道徳ジレンマに対する判断をしてもらった後に、「なぜそのように判断したのか」と尋ねると、正当化の理由を適切に述べられなかった（あるいは全く述べることができなかった）人が大多数であった。感情説では、これは道徳判断を生み出すのは我々の無意識的・自動的・直観的な感情であるということを示す証拠だと考える。これに対して、道徳と言語のアナロジー説では、これは道徳判断を生み出すのは無意識的に作用する直観的な認知能力としての道徳文法だということを示す証拠だと考える。だが、人々は道徳判断を意識的な推論によって行っているというピアジェ＝コールバーグ流の考え方では、大多数の人間が正当化の理由を述べられないのに道徳判断をすることができるという事実を説明することができない。

両者の第二の共通点は、道徳判断には一定の普遍性がみられるという主張である。これは既に人類学によってある程度明らかにされた事実であると言えるが、ハウザーらはこれを定量的に示そうという試みを実施している。彼らは「道徳感覚テスト (the Moral Sense Test)」という実験をインターネットで公開し、世界中からアクセスした人々に道徳的ジレンマについて回答してもらった。その結果は、文化・社会的背景に関わらず一定であった (Hauser et al. 2008)。彼らによれば、こうした道徳判断の普遍性は、普遍道徳文法が存在を支持する結果ということになる。他方でハイトやグリーンの感情説によれば、道徳判断の普遍性は生得的な感情によるものと判断される。

次に、両者の相違点について指摘しよう。第一に、道徳と言語のアナロジー説は、原理・パラメータ・モデルを採用していることから明らかなように、道徳判断の普遍性だけでなく、文化ごとの多様性も強調している点で、ハイトやグリーンの感情説とは異なる (プリンツの感情説に関しては4節を参照)。道徳の文化間の多様性は確かに存在するので、これを視野にしっかりと入れている点で道徳と言語のアナロジー説は高く評価することができるだろう。だが現在のところでは、原理については候補 (二重結果の原理) が挙がっているものの、そのパラメータについては何も言及されていない。したがって、目の付け所は評価できるとは言え、絵に描いた餅に過ぎないということ是否めないだろう。

第二に、道徳判断を生み出す心理過程の性質に関して、両者の見解は異なっている。感情説によれば、道徳判断を生み出すのは主として感情である。彼らは感情という言葉で認知的な推論過程とは全く性質の異なる心理過程を意味している。これに対して、

道徳と言語のアナロジー説は道徳判断を生み出すのはあくまでも認知計算過程（ただし、これは無意識的なものである）だと主張する。そして、この認知計算過程の性質を具体的に明らかにすることが、重要な研究課題であると考えている。その結果提出されたのが、行為分析モデルである。したがって感情説と道徳と言語のアナロジー説では、道徳判断の生成に関して「感情」と「行為分析」という異なる心理過程が強調されているのである。

§4 道徳と言語のアナロジー説への批判

前節では道徳と言語のアナロジー説を分析し、感情説との比較を通じてその特徴をさらに明確にした。道徳と言語のアナロジー説は、確かにいくつかの興味深い論点を提供しているし、その点では評価できる。とはいえ、この説にはいくつかの批判がなされてきている。本節ではこの批判を検討する。

道徳と言語のアナロジー説への批判者として第一に挙げられるのは、プリンツ（Prinz 2008a, 2008b, 2008c）であろう。彼はまず、言語と道徳との間に見られるいくつかの大きな相違点を指摘する（Prinz 2008b）。第一に、言語の習得には極めて重要な時期が存在し、それ以後の言語習得は困難となるが、道徳の場合は成人後であっても比較的容易に異なる道徳規範体系を身に付けることができる。第二に、道徳習得の場合には、親のしつけなどにおける否定的もしくは肯定的なフィードバックが大きな役割を果たしている。すなわち、子どもが道徳を身に付けるためには、規範を犯したときに叱られ、遵守したときには褒められるというプロセスが重要なのである。しかし、言語習得においては、道徳の場合ほど上記のようなフィードバックは大きな役割を果たさない。第三に、様々な道徳規範の中には互いに対立するものが存在し、いずれが正しいのかということが重要な問題となる。他方で、述語を目的語の前に置くことと後に置くことのいずれが正しいのかについて深刻な論争が生じることはないだろう。

以上の論点は、言語と道徳のアナロジーを弱めるものであっても、否定するものではないのかもしれない。しかし、プリンツは道徳と言語がさらに根本的な点で異なっていると主張する。まず、言語の場合、特定の脳領域を損傷すると、その他の能力には問題がないにも関わらず、言語に関する能力だけが損なわれるという事例（例えば様々なタイプの失語症など）が知られている。道徳の場合も同様に、その他の能力から独立した能力であると考えられるかもしれない。しかし、道徳が欠如していると言われるサイコパスや前頭前皮質の一部を損傷した患者は、同時に感情に関わる能力（表

情から他者の感情を判断する能力など)も欠損しているということが知られている。したがって現在のところ、道徳が他の能力から独立した能力であるということをはっきりと示す証拠はない。次に、言語使用と道徳判断の背後には無意識的な文法が存在する、という点にもプリンツは懐疑的である。既に述べたように、道徳と言語のアナロジー説の支持者は、道徳的ジレンマに対して道徳判断を下す実験の中で被験者が自身の判断を上手く正当化できないことから、道徳判断は無意識的な道徳文法によって導かれているのであり、言語の場合と同様に、文法に対して意識的にアクセスすることはできないのだと主張していた。しかし、「人を殺すな」というような単純な道徳規則に関しては、容易に意識的なアクセスが可能であるように思われる。

さらに、彼は言語とのアナロジー説支持者が根拠の一つとして挙げている慣習と道徳の区別についても反論を試みている(Prinz 2008b)。たとえば一部の心理学者は、我々は経験が極めて少なくても、慣習と道徳を区別できるようになると主張している(Nucci 2001 など)。彼らによれば、道徳は(1)慣習よりも深刻で、(2)危害に関する言及を含み、(3)何の権威にも依存しない客観的な規範である(Turiel 1998)。しかし、プリンツによれば、第一に、道徳規範に対する違反(分け合うことを提案せずに最後のクッキーを食べてしまうこと)よりも慣習に対する違反(裸で出勤すること)の方が深刻な場合がある。第二に、危害に関する言及を含む慣習(腐った肉を食べてはならない)がある一方で、そうした言及を全く含まない規範(国旗をトイレに流すべきでない)が道徳的であると見なされる場合もある。第三に、宗教に基づく規範は権威(すなわち神)に依存するが、多くの人はこれを道徳規範と見なしている。他方で、慣習であっても権威に依存しない規範(腐った肉を食べてはならない)も存在する。以上のような議論は、慣習/道徳の区別そのものを否定するわけではないが、両者を区別する上記の基準が正しくないことを示している。そうすると、慣習と道徳の区別が生得的能力の産物であるという主張も説得力を失うことになる。

また、道徳判断の多様性を原理とパラメータで説明しようという点にも大きな問題がある(e.g., Dupoux and Jacob 2007; Prinz 2008b)。たとえば危害を加えることが許容される範囲について、一体どのようなパラメータが想定できるだろうか。まずは、危害を加える対象に関するパラメータが考えられる。たとえば、一親等かそうでないか、二親等かそうでないか、あるいは非血縁者か、などと、これだけでもパラメータは無数に必要なかもしれない。さらに、危害の種類についてもパラメータが存在するかもしれない。殴ってもよいか、コミュニティから追放してもよいか、殺してもよいか...と考えていくと、こちらも無数のパラメータが要求されてしまうだろう。もちろん、

どの程度のパラメータが要求されるのかはやってみなければ分からないが、言語のように二値のパラメータで説明しようとするなら、おそらく相当な数のパラメータが要求されてしまうように思われる。

以上のような批判に基づき、プリンツ (Prinz 2008a, 2008b) は感情説の立場から多様性の説明を試みている。プリンツによれば、道徳は共感、模倣、心の理論などの道徳以外の領域でも役割を果たすいくつかの心理能力によって学習されていく。例えば、危害原則の学習について考えてみよう。危害を加えられた他者が抱く不快感や嫌悪感、共感を通じて我々のうちに再現される。そして、心の中に相手と同様の感情が再現され、さらに心の理論によって他者の立場に立つことができれば、「他者に危害を加えるべきでない」という規範が学習されるはずだという (詳細は Prinz 2008a を参照)。ここで、プリンツは共感や不快感、嫌悪感それ自体を生得的なものと見なす一方、何に対して不快感を抱くかについては学習されるものだと考えている。例えば、非常に生存競争が厳しくて非血縁者の死を喜ぶような文化で育てば、生得的な快感や共感を通じて「血縁者に対しては危害を加えるべきではない」という危害原則が学習されるかもしれない。逆に何人たりとも殺してはいけないという文化で育てば、「いかなる場合にも他人に危害は加えるべきではない」という原則を学習するかもしれない。こういった学習過程によって道徳判断の多様性は説明できるというのが彼の主張である。

これらの批判に対して、ハウザーらは発見法的役割が果たせれば十分だと考えているようだ (Hauser, Young, and Cushman 2008; Dwyer and Hauser 2008)。確かに、行為分析モデルなどは言語とのアナロジーに立ってはじめて導出が可能であったものであり、この点に関して言語とのアナロジー説の功績は否定しがたいものがある。とはいえ、道徳と言語の間には上記のような無視できないギャップがあるため、その発見法的役割もまた、非常に限定的なものにならざるを得ないようにも思われる。

§5 今後の展望

言語とのアナロジー説が抱えるもう一つの課題として、感情の取り扱いを挙げることができだろう。ハウザーらの考えでは、無意識的な行為分析の後に道徳判断が下され、感情は判断の後に生み出されるという主張がなされている (e.g., Hauser 2006)。しかし、これはあまりにも感情説の主張を軽視しすぎた議論であろう。感情説では逆に、無意識的な感情プロセスが道徳判断の背後にあると論じてきたのであり、2節でも論じたように、このような説明は一定の成功を収めてきているのである。

もちろん、道徳と言語のアナロジー説に立つ論者が感情を重視しないことには理由がある。たとえば、グリーンの個人的／非個人的な状況の区別はトロリー・ジレンマと歩道橋ジレンマに対する反応に対してある程度の説得力を持っていたが、ミハイルやハウザーらが扱ったジレンマ（泣き叫ぶ赤ん坊のジレンマと嬰兒殺しのジレンマ）では、この区別で説明ができないような回答の変化が見られている（Mikhail 2007, p. 149, table 1）。非個人的なジレンマでも少し状況（傷害の数や、傷害が副産物であるか否か、など）を変えてやれば、人々の道徳判断の内容は変化するのである。彼らは自身の行為分析モデルに基づいて、この変化を傷害の数や傷害が手段か副産物かという基準により見事に説明できると考えている（ibid., p. 150）。彼らの説明は確かに上手くっており、行為分析モデルは一定の説得力を持つと言える。また、感情説ではこのような認知計算過程に関する考察はほとんど行われてこなかった。その意味でも、行為分析モデルは一定の価値を持っていると言えるだろう。

とはいえ、行為分析モデルがジレンマに対する説明を上手く行えたとしても、感情の入り込む余地が全くないわけではない。たとえば、ミハイルの分析によれば傷害の数に応じて許せる度合いが変化しており、これは義務論規則によるものだと考えられている。しかし、これは傷害の数に応じてより強い否定的感情を誘発しているだけかもしれない。ここでの傷害の数は無意識的に判断されるものであるが、単純な例で考えるなら、殴られた数だけ強い否定的感情を抱くということは十分に考えられる。さらに、傷害が副産物であるかどうかについても、傷害が持つ重要性が変化しているだけであるのかもしれない。自身の行為の直接的結果と副産物では、前者により強い感情的反応を示すということもありうる。

以上のように、道徳と言語のアナロジー説を支持するにしろ感情説を支持するにしろ、行為分析モデルによる説明と感情による説明には折り合いをつけねばならないだろう。というのも、感情が道徳判断の生起に全く関わっていないということは神経科学の実験結果からもありえない話であるし、他方で感情を説明の中心に据えるだけで満足し、行為分析モデルのような精緻な道徳判断生起モデルを作らないままでは、ハウザーらが明らかにした様々な道徳的ジレンマに対する人々の道徳判断の違いを説明できないからである。このように現時点では、道徳と言語のアナロジー説と感情説の両者に改善の余地があるように思われる。したがって両者の調停は、道徳と言語のアナロジー説が感情概念を適切に取り込んでいくことができるかどうか、そしてまた感情説がより詳細な道徳判断生起モデルを作っていくことができるかどうかにかかって

いると言えるだろう⁶。

参考文献

- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: The MIT Press. [邦訳『文法理論の諸相』, 安井稔訳, 研究社, 1970年.]
- Damasio, Antonio R. 1994. *Descartes' error: Emotion, reason, and the human brain*. New York: Harper Perennial. [邦訳『生存する脳: 心と身体の神秘』田中三彦訳, 講談社, 2000年.]
- Dupoux, Emmanuel and Pierre Jacob. 2007. Universal moral grammar: A critical appraisal. *Trends in Cognitive Sciences* 11: 373–378.
- . 2008. Response to Dwyer and Hauser: Sounding the retreat? *Trends in Cognitive Sciences* 12: 2–3.
- Dwyer, Susan. 1999. Moral competence. In *Philosophy and linguistics*, ed. K. Mura-sugi and R. Stainton, pp. 169–190. Boulder, Colorado: Westview Press.
- . 2006. How good is the linguistic analogy? In *The innate mind, volume 2: Culture and cognition*, ed. P. Carruthers, S. Laurence, and S. Stich, pp. 237–256. New York: Oxford University Press.
- Dwyer, Susan and Marc Hauser. 2008. Dupoux and Jacob's moral instincts: Throwing out the baby, the bathwater and the bathtub. *Trends in Cognitive Sciences* 12: 1–2.
- Frank, Robert H. 1988. *Passions within reason: The strategic role of the emotions*. New York: Norton. [邦訳『オデッセウスの鎖: 適応プログラムとしての感情』, 山岸俊男監訳, サイエンス社, 1995年.]
- Greene, Joshua. 2005. Cognitive neuroscience and the structure of the moral mind. In *The innate mind, volume 1: Structure and contents*, ed. P. Carruthers, S. Laurence, and S. Stich, pp. 237–256. New York: Oxford University Press.
- Greene, Joshua, Leigh Nystrom, Andrew Engell, John Darley, and Jonathan Cohen. 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44: 389–400.

⁶ 初期段階の杜撰な草稿に対して, 東京大学大学院人文社会系研究科の矢島壮平氏, 京都大学大学院文学研究科の伊勢田哲治氏, 杉本俊介氏から貴重なコメントを頂いた。また, 2名の査読者からも極めて的確かつ有益なコメントを頂いた。ここに記して感謝したい。

- Greene, Joshua, Brian Sommerville, Leigh Nystrom, John Darley, and Jonathan Cohen. 2001. An fMRI investigation of emotional engagement in moral judgement. *Science* 293: 2105–2108.
- Haidt, Jonathan. 2001. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review* 108: 814–834.
- Harman, Gilbert. 2000. Moral philosophy and linguistics. In *Explaining value: And other essays in moral philosophy*, ed. G. Harman, pp. 217–226. New York: Oxford University Press.
- Hauser, Marc. 2006. *Moral minds: How nature designed our universal sense of right and wrong*. New York: Ecco Press/HarperCollins.
- Hauser, Marc, Fiery Cushman, Liane Young, R. Kang-Xing Jin, and John Mikhail. 2007. A dissociation between moral judgment and justifications. *Mind & Language* 22: 1–21.
- Hauser, Marc, Liane Young, and Fiery Cushman. 2008. Reviving Rawls's linguistic analogy: Operative principles and the causal structure of moral actions. In *Moral psychology, volume 2: The cognitive science of morality: Intuition and diversity*, ed. W. Sinnott-Armstrong, pp. 197–243. Cambridge, Mass.: The MIT Press.
- Mikhail, John. 2000. Rawls' linguistic analogy: A study of the 'generative grammar' model of moral theory described by John Rawls in 'A theory of justice'. PhD diss., Cornell University.
- Mikhail, John. 2007. Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences* 11: 143–152.
- Nucci, Larry P. 1965. *Education in the moral domain*. New York: Cambridge University Press.
- Prinz, Jesse. 2008a. *The emotional construction of morals*. New York: Oxford University Press.
- . 2008b. Is morality innate? In *Moral psychology, volume 1: The evolution of morality: Adaptations and innateness*, ed. W. Sinnott-Armstrong, pp. 367–406. Cambridge, Mass.: The MIT Press.
- . 2008c. Resisting the linguistic analogy: A commentary on Hauser, Young, and Cushman. In *Moral psychology, volume 2: The cognitive science of morality: Intuition and diversity*, ed. W. Sinnott-Armstrong, pp. 157–170. Cambridge, Mass.:

The MIT Press.

Turiel, Elliot. 1998. The development of morality. In *Handbook of child psychology, vol. 3: Social, emotional, and personality development*, ed. N. Eisenberg, pp. 863–932. New York: Wiley.

田窪行則・中島平三・福井直樹・稲田俊明・外池滋生．1998 年．『岩波講座 言語の科学（6）生成文法』東京：岩波書店．

